

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2002, 22:41:46 ; Search time 2691.5 Seconds
(without alignments)
16489.608 Million cell updates/sec

Title: US-10-025-514-15
Perfect score: 1525
Sequence: 1 tcttagaccatggaaagacct.....ccagtcaaggccatgtcgac 1525

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenBambl:
1: gb_ba:
2: gb_htg:
3: gb_ln:
4: gb_on:
5: gb_ov:
6: gb_dat:
7: gb_bh:
8: gb_pl:
9: gb_pr:
10: gb_ro:
11: gb_sts:
12: gb_sy:
13: gb_jn:
14: gb_vl:
15: em_ba:
16: em_fun:
17: em_hum:
18: em_in:
19: em_mu:
20: em_om:
21: em_or:
22: em_ov:
23: em_pat:
24: em_ph:
25: em_pl:
26: em_ro:
27: em_sts:
28: em_uni:
29: em_vl:
30: em_htg_hum:
31: em_htg_inv:
32: em_htg_other:
33: em_htg_mus:
34: em_htg_pln:
35: em_htg_rcd:
36: em_htg_mrn:
37: em_htg_vrt:
38: em_sy:
39: em_htgo_hum:
40: em_htgo_mus:
41: em_htgo_other:
Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match %	Length	DB ID	Description
1	629.4	41.3	13.08	6	AR111412	Sequence
2	628.4	41.2	11.85	6	AR111411	Sequence
3	437	28.7	13.45	9	HUMATB	MI1465 Human alpha
4	433.8	28.4	14.34	6	E00631	E00631 DNA encodin
5	432.2	28.3	13.12	6	I02706	I02706 Sequence 1
6	430.6	28.2	15.84	9	BC011991	Homo sapi
7	429	28.1	11.85	6	AR111410	Sequence
8	429	28.1	13.52	6	AX35339	Sequence
9	429	28.1	13.71	6	AX35338	K01396 Human lpha
10	429	28.1	13.71	9	HSATP41	AX35338 Sequence
11	429	28.1	13.71	9	X01683	Human mRNA
12	429	28.1	13.99	9	AK026174	AK026174 Homo sapi
13	429	28.1	14.34	6	E00195	E00195 cDNA encodin
14	429	28.1	14.34	6	I04196	I04196 Sequence 3
15	429	28.1	14.34	6	I04272	I04272 Sequence 1
16	429	28.1	14.34	6	I07849	I07849 Sequence 2
17	429	28.1	14.35	6	AX019486	AX019486 Sequence
18	429	28.1	14.35	6	Af130068	Af130068 Homo sapi
19	427.4	28.0	13.37	9	J02619	J02619 Human Z typ
20	427.4	28.0	13.78	6	BC015642	BC015642 Homo sapi
21	427.4	28.0	14.31	9	Af113676	Af113676 Homo sapi
22	427.4	28.0	14.31	9	I02998	I02998 Sequence 8
23	427.4	28.0	12.99	6	I07949	I07949 Sequence 2
24	425.8	27.9	12.99	6	BABILATA	J00321 Papilio hamad
25	425.8	27.9	13.08	9	I00556	I00556 Sequence 2
26	425.8	27.9	13.78	6	I03509	I03509 Sequence 6
27	425.8	27.9	13.78	6	I07947	I07947 Sequence 4
28	425.8	27.9	14.34	6	I01352	I01352 Sequence 1
29	425.4	27.9	11.85	6	A01846	A01846 Human mRNA
30	424.2	27.8	13.78	6	I01227	I01227 Sequence 2
31	417	27.3	11.91	9	AB004044	AB004044 Cercopith
32	407	26.7	13.90	6	AX202089	AX202089 Sequence
33	404.6	26.5	13.56	6	I36163	I36163 Sequence 16
34	403	26.4	13.56	6	I36161	I36161 Sequence 12
35	402.6	26.4	20.13	6	AX472008	AX472008 Sequence
36	399.8	26.2	13.56	6	I36164	I36164 Sequence 18
37	390.2	25.6	13.51	10	AB000550	AB000550 Spemophil
38	387	25.4	13.72	10	AB000552	AB000552 Callosciu
39	385.4	25.3	13.51	4	X88780	X88780 S.scrofa mR
40	378	24.8	13.06	10	M32247	M32247 Rat alpha-1
41	377.2	24.7	13.52	10	S77822	S77822 alpha-1-antitrypsin
42	376.4	24.7	13.80	10	D00675	D00675 Rat mRNA
43	372.8	24.4	13.43	10	AB000546	AB000546 Tamias ssi
44	372.6	24.4	12.42	6	E13268	E13268 Tamias ssi
45	370.6	24.3	13.80	4	X63129	X63129 B.taurus mR

ALIGNMENTS

RESULT 1	AR111412	1308 bp	DNA	linear	PAT 14-FEB-2001
LOCUS	Sequence 6 from Patent US 6127145				
DEFINITION					
ACCESSION	AR111412				
VERSION					
KEYWORDS					
SOURCE					
ORGANISM	Unknown				
UNCLASSIFIED	Unclassified				
REFERENCE	1 (bases 1 to 1308)				
AUTHORS	Sutliff, T.D. and Rodriguez, R.L.				
TITLE	Production of alpha-1-antitrypsin in plants				
JOURNAL	Patent: US 6127145-A 6 Oct-2000;				
FEATURES	Location/Qualifiers				

Qy	492	GTTGAGAAAGGACCCAGGGTAAGATCGTTGACCTAGTTAAAGAAATTAGATCGTGATACC	551		FEATURES	source	MEDLINE PUBMED COMMENT	89198065 2559391
Db	481	GTCGAGAAAGGACCCAGGGCAAGTGTGCAAGGATTCAGGGAAACCCTTCGAGGT	540				[21] sites; alpha-1-antitrypsin deficiency disease mutations. The Granite Falls (a one base pair deletion), Bellingham (a substitution) and Mattawa (a one base pair insertion) mutations all result in premature transcription termination.	
Qy	552	GTCTTCCACTAGTAACTATTTTCAAGGGTAAGTGGAAACCCCTTCGAGGT	611					
Db	541	GTCCTGGCTGTCACATCTTCAGGCAACTGGCAAGGGCAACTGGCAAGTGGTG	600					
Qy	612	AAAGATACTGAAAGGAAAGATTTCATCTGTGATAAAGTTACTGTAAGTCCAACTG	671					
Db	601	AGGACACCGAGGAGGACTTCCACGTCACAGGCAACCCGTCAGGTGGATG	660					
Qy	672	ATGAAAAGACTGGPATGTCATATTACATGCAACATGCAAAATAAAGTCTTGGGTATA	731					
Db	661	ATGAAAGGCTGGATGTCAACTGCAAGCTGGAGGTCTCAGGTGGTCTC	720					
Qy	732	TTAATGAAAGTATTAGGTAACGCTACTGTTATTTCAGGAGCAAGTAACTT	791					
Db	721	CTCATGAAAGTACCTGGGAACGGCACCACGCCATCTTCTGGGAGGGCAAGCTC	780					
Qy	792	CAACATTAGAAATGAGTTGACTCATGACATTATTACTAAATTTTAGAGAACGAGGAT	851					
Db	781	CAGCACCTGAGAAAGGAGCTGACCCACATCATAGAACTTCCCTGGAGACGAGGAC	840					
Qy	852	CCTGCTTAGGCTCTCTGACCTGGCAAAAGTAAAGTACCGGTACTAGCCTAAAA	911					
Db	841	AGGCCCTCCCTGCTCCACCTGGCAAGGCTGAGCATCACGGCAACTGGAGAAG	900					
Qy	912	TCTGTTTAGGCCAGTTAGGTATTACCAAGTTTCTTAAGGGCCGATTGACTGGT	971					
Db	901	AGGCGTGTGGCCACGGCTGGCATCACGAGGCTCTAGGCAAGCTCCGGC	960					
Qy	972	GTTACTGAGAAAGTCCATTAAATTGGTAAGCTGTCTCACAAAGCGCTTAACTATT	1031					
Db	961	GTGACGGAGGAGGCCCTGAGCTCTCCANGCCCTGCAAGCGGTGCTCAGARC	1020					
Qy	1032	GATGAAAGGTACCGGAGCCGCTATGGCTGGAAAGTATTCCAATGAGCATT	1091					
Db	1021	GACGAGAGGGGGAGGAGCTGGGGCCATGCCACGGGCACTGGCCATCCATG	1080					
Qy	1092	CCACCGAGGTAAATTAAACCATTCGTTCTGATGATCGAGGAGAACACTAA	1151					
Db	1081	CCGCCGAGGTCAAGTCACAGGCCCTCGCTCTCCGTGATGCCAGAACCGAAG	1140					
Qy	1152	AGGCCATGTTTATGGTAAAGGTTCTCAACCCAACTCAGANG	1193					
Db	1141	AGCCCCCTCTCATGGGAAGGTCTGTCACCCACCCGAGAAG	1182					
RESULT 3								
LOCUS	HUMAN1TB	1345 bp	mRNA	linear	PRI 30-OCT-1994			
DEFINITION	Human alpha-1-antitrypsin	in mRNA, complete cds.						
ACCESSION	M11465							
VERSION	M11465.1	GI:177826						
KEYWORDS	alpha-1-antitrypsin.							
SOURCE	Human liver. CDNA to mRNA.							
ORGANISM	Homo sapiens							
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.								
Curie, D., Brantly, M., Curiel, E., Ster, L., and Crystal, R.G.								
2 (sites)	Cilliberto, G., Dente, L., and Cortese, R.							
Cell-specific expression of a transfected human alpha 1-antitrypsin								
JOURNAL	Cell 41 (2), 531-540 (1985)							
MEDLINE	85175977							
PUBMED	2985281							
REFERENCE	2 (sites)							
AUTHORS	Curie, D., Brantly, M., Curiel, E., Ster, L., and Crystal, R.G.							
TITLE	Alpha 1-antitrypsin deficiency caused by the alpha 1-antitrypsin							
Nullmattawa gene. An insertion mutation rendering the alpha 1-antitrypsin								
JOURNAL	J. Clin. Invest. 83 (4), 1144-1152 (1989)							
Qy	372	TCTGAGGTAAATTGCTGTTACCGTTACGGTAAATTGCTGAAACTATATCAGT	431					
Db	444	AGCGAGGGCTGAAACTAGGGATAAGGTTGGAGGTAAAAGTGTACCACTCA	503					
Qy	432	GAGGCCTTACCGTTACGGTAAATTGGTACTGAGGAAGTAAAGCAAAATTATGATT	491					

Query Match 28.3%; Score 432.2; DB 6; Length 1312;
 Best Local Similarity 60.4%; Pred. No. 5; 5e-82; Indels 0; Mismatches 466; Caps 0;

RESULT 5
 102706 LOCUS 102706 1312 bp ss-DNA linear PAT 21-MAY-1993
 DEFINITION Sequence 1 from Patent US 4599311.
 ACCESSION 102706.1 GI:268359
 VERSION 1
 KEYWORDS Unknown.
 SOURCE Unknown.
 ORGANISM Unclassified.
 1 (bases 1 to 1312)
 REFERENCE Kawasaki, G.H.
 AUTHORS Glycolytic promoters for regulated protein expression: protease
 TITLE Inhibitor
 JOURNAL Patent: US 4599311-A 1 08-JUL-1986;
 1547 - 16th Ave. East, Seattle, WA
 Location/Qualifiers 1..1312
 FEATURES /organism="unknown"
 source BASE COUNT 339 a /organism="unknown"
 ORIGIN 368 c 324 g 281 t

Query Match 28.3%; Score 432.2; DB 6; Length 1312;
 Best Local Similarity 60.4%; Pred. No. 5; 5e-82; Indels 0; Mismatches 466; Caps 0;

QY 599 GTGGAAAGGTACTCAAGGAAAATTGTGGATTGGTCAAGGAGCTTGACAGACACA 658
 DB 552 GTCCTTCGCACTTAACCTATTTCAGGTAAGTGGAACTTTCGAGGT 611
 QY 659 GTTTTGTCTGTGAATTACATCTCTTAAAGGCAAATGGAGAACCTTGAACTG 718
 DB 612 AAAGATACTGAAGGGAAATTCTGATCAAGTACTCTGCAAGTCCAATG 671
 QY 719 AAGGACACCGGAAAGGAGACTCCAGTGGACCCAGGTGAAAGTGGCTAIG 778
 QY 672 ATGAAAGACTGGTATGTCATAATTCAATGCAAAATTAAAGTCTGGTCATA 731
 DB 779 ATGAAAGCTTAAAGCTGTTAACATCAGACTGTAAGAGCTGTGAGCTGGCG 838
 QY 732 TTATGAGTATTAGTAACTCTGTTACCTGTTACCTGAAAGGTAAGCT 791
 DB 839 CTCATGAATACTGGCAATGCCAACCGCCATCTCTCTGTGATAGGGAAACTA 898
 QY 792 CAACATTAGAAATGAGTTGACTCATGACATTAACTAAATTAGAAACGAGAT 851
 DB 839 CACACCTGGAAATGAACTACCCAGATACTACCAAGTTCCTGGAAATGAGAC 958
 QY 852 CGTCGTAGCGTTCTGCACTGCAACTAACTAACCTGGTACTTACGACTTAA 911
 DB 959 AGAAGGTC TGCCAGCTTACATTACCCAAACTGTCCATRACTGGAACCTATGATCTGAG 1018
 QY 912 TCGTTTATGGCAGTGTAGGATTACCAAGTTCCTGCACTGCAAGTGGCT 971
 DB 1019 AGCGTCTGGTCACATGGGATCACTAAGCTGGTCACTAAGTGTGACCC 1078
 QY 972 GTTACTGAGAAGCTCATTAAATTGTGAAAGCTGTTACAAAGCCCTTAACATT 1031
 DB 1079 GTACAGAGGGGACCCCGCCGCGCCGCGCATTTAGGGCTCATAGGCT 1138
 QY 1032 GATGAAAGGGTACCGAGCGCCGGCGCATGTTCTGAACTGATGAGCT 1091
 DB 1139 GACGAGAAAGGACTGAAGGACTGTCGGCCCATGTTTAAAGGCCATCCATGCTATC 1198
 QY 1092 CCACCAAGGTTAAATTATAAACCATATGTTCTGATGTCAGCAGACACTAA 1151
 DB 1199 CCCCGCAGGTCAAGTCAACAAACCTTGTCTCTATGATGTTGAAACAAATAACCAAG 1258
 QY 1152 AGCCATTTGTTATGGTAAAGGTTGCAACCCAACTCAGAA 1192
 DB 1259 TCTCCCTGTGTCATGGAAAGTGTGATCCACCCAAA 1299

QY 12 GAAGACCCCTCAAGGCCAACGCCGCTAAAAAACCGCACCCAGTCAACGACAGACAT 71
 DB 100 GAGGATCCAGGAGATGCTGCCCAGGAGACATCCCAACCTGATCGATCAC 159
 QY 72 CGACCTTTAATAAAATTACTCCAAATTAGCCGAATTTCGCTTTTCTTGTATAGC 131
 DB 160 CCAACCTTCACAAAGTCAACCCCAACTCTGGCTGAGTTCAGCTTACGGAG 219
 QY 132 TTAGCTCATCAAGTAACTCTACTACATTTCAGTGGCTCTTATTGCCAGCT 191
 DB 220 CTGGCACACCGTCACAGCTGAGTCACTGCTCACGCCAGTACCC 279
 QY 192 TTGCGCATGTGAGTTAGSTGACTAAAGCGATACCATGACGAGATTAGAGGTTA 251
 DB 280 TTGCAATGGCTCCTGGGACCAAGGGTGCACACTACGATGAAATCTGGAGGGCTG 339
 QY 252 AACCTTAATTGACCCAAATCCCAAGGCCAAATTCAAGGGGTTTCAGAGTTTG 311
 DB 340 AATTGCAACCTCACGGAGATTCCGGAGTCAGTCATGAAGGCTTCAGAACCTCTC 399
 QY 372 TCTGAGGTTAAATGTGTTGACAAATTCTGAGTCAGAAACTATATCATAGT 431
 DB 460 AGCGAGGGCTGAACCTAGGGATAAGTTGGAGGTGTTAAAGGTTGACCTCA 519
 QY 432 GAGGCCTTTACCGTTATTGGTGTACTGAGGAAGTAAAGAAATTATGATATT 491
 DB 520 GAAGGCTTCACGTGCACTTCGCGACACCGGAAGGCCAAACGATCAACGATAC 579
 QY 492 GTTGAAGAAAGGCCAACCGGGTACATGGTGTGACCTGTTAAAGAATTTAGATCTGATACC 551
 DB 580 GTGGGAGAAGGTTACTCAAGGAAAATTGTGGATTGTGTCAGAGGCTGTACAGAGACA 639
 QY 552 GTCTGCGACTTAGTTAACTATATTTCAGTGGAAACTCCPTTCGAGGTT 611
 DB 640 GTTTTGCCTGGTGAATTACCTCTTTAAAGGAAATGGAGACACCCCTTGAGTC 699
 DB 760 ATGAGACGGTTAGGSCATTACATCAGCATTGAGACGCTGTCAGCTGGCTCTG 819
 QY 732 TTAATGAAAGTATTGGTAACGGTACTGCTATTAACTGAAAGGTAAGCCTTA 731
 DB 820 CTGATGAAATACCTGGCAATGCCCATCTCTCCGCTATGAGGGAAACTA 879
 QY 792 CAACATTAGAGAATGAGTTGACTCATGACATTATTAACTAAATTAGAAGC 791
 DB 880 CAGCACCTGAAATGAACCTACCCACCATCATACCAAGTCTGGAAAATGAAGAC 939
 QY 852 CGTGTAGGGCTTCCTCAGCTGGCAAGTAAGTTCACCGGTACTTACGACTTAA 911
 DB 940 AGAGGTCTGCAAGCTTACATTCACCAAACTGTCAATTACTGGAACTATGATCTGAG 999
 QY 912 TCGTTTAGGCCAGTAGGTTTCAACGGTGCAGATTGAGTGGT 971
 DB 1000 AGCCTCCTAGGTCAACTGGCATCACTAAGGCTTTCAGCTTCCGG 1059
 QY 972 GTTACTGAGAAGGCTCATTAAATTGGATAAGCTTCAACGGCTCTTAACATT 1031
 DB 1060 GTCAAGGAGGAGGACCCCTGAGCTCCTGGCCATGAGCTGCTGACCATC 1119
 QY 1032 GATGAAAGGGTACCGGGCCGGCTATGTTCTGACCAAAAGCTTCAACGGCTCTCCGG 1091
 DB 1120 GACGAGAAAGGACTGAGCTGGCCATGAGCTGGCCATGAGCTGCTGACCATC 1179

Qy	1092	CCACAGAAGTTAAATTAACCATTGTTCTGATGATCGAGAACACTAA	1.151	BASE COUNT	410	a	459	c	379	g	336	t
Db	1180	CGCCCGAGCTCAAGCTTCACAAACCCCTTGCTCTTAATGAACTAACAG	1.239	ORIGIN								
Qy	1152	AGCCCATTTGTTATGGTAAGTGTCAACCACACTAGAA	1.192	Query	28.2%	Match	28.2%	Score	430.6	DB	9	Length 1584;
Db	1240	TCTCCCTCTCATGGAAAGTGTGAATCCCACCCAAA	1.280	Best Local Similarity	60.3%	Pred.	No. 1.2e-81;	Mismatches	469	Indels	0	Gaps 0;
RESULT	6			Matches	712	Conservative	0					
BC011991				Qy	1.2	GAAGACCCCTCAAGGAGGCCCTCAAAACCGAACACTCATCACGACCAAGACCAT						
LOCUS				Db	305	GAGATCCCGAGGAGATGCTGCCAGAGACATACTCCACCATGATCAGGATCAC						
DEFINITION				Qy	7.2	CCGACTTTAATAAAATTACTCCAAATTAGCGAAATTGCTTTCCTGTATAGACAA	1.31					
MGC-9222				Db	365	CCACCCCTAACAGATACCCCAACCTGGCTGAGTTGGCTACAGCTACGCC						
IMAGE:3859644				Qy	13.2	TAGCTCTCAAACTAACTAATCTACTAACTAATTTAGCTGTTCTATGCCACTGT	1.91					
BC011991				Db	425	CTGGCACACGCCAACGCCCCATCTCTCCCAAGTGGCATGCCTACGCC						
1				Qy	19.2	TTCGCCAGTTGAGTTAGTTAGCTAACCGATACCATGAGGAGTTAAAGGTTA	2.51					
GI:15080498				Db	485	TTCGCAATGCTCTCCCTGGAGAACCTGACACTACGAAATCTGGCGCTG						
MGC				Qy	25.2	AACPTTAATTGACCCATTAACTGCAATGAGGTTCAGAGTTGTT	3.11					
ORGANISM				Db	545	ATTTCAACCTCACGGGATTCGGGCTCATGGCTACAGCTACAACTCC						
Homo sapiens				Qy	31.2	AGACCTTAACTAACCTGATTCATGAAATTGCTAACCTGTAAGGTTATTTG	3.71					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				Db	605	CGTACCCCTAACCCGGCTCCAGCTGACACCCGAATGGCTGTC						
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				Qy	372	TCTGAAGGTTAAATTGCTGAAATTCCCTAGAAAGCTTATATCATG	4.31					
1 (bases 1 to 1584)				Db	665	AGCGAGGCGCTGAACTGGTAAAGCTGGTTGAAAGCTTGGTAAAGCTCA						
Strausberg, R.				Qy	43.2	GAGGCTTAACTCCCTGGTAACTGGTAACTGGTAACTGGTAACTGGT	4.91					
Direct Submission				Db	725	GAAGGCTTCACTGTCACATTGGGACACCCGGTCCACGTTGAGCT						
Submitted (30-JUL-2001) National Institutes of Health, Mammalian				Qy	552	GTCTTCGACTAGTTACCTATATTTCAAGGTAATGTTGACCTTCA						
Gene Collection (MGC) National Genomics Office, National Cancer				Db	785	GTGGAAAGGCTGAACTGGTAACTGGTAACTGGTAACTGGT						
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,				Qy	592	GTGGAAGGCCCCAGGTAACTGGTAACTGGTAACTGGTAACTGGT						
USA				Db	845	GTTCCTGCTCTGGGTAATGACCTTCACTGGGAAACCTTGGTAA						
NIH-MGC Project URL: http://mgc.nci.nih.gov				Qy	612	AAAGATACTGAGGAGATTCATGGTAACTGGTAACTGGTAACTGGT						
Contact: MGC help desk				Db	905	AGGAGACCCGGGAGGAGGACTCCACGTTGAGGTGAACTGGGAAACTA						
Email: cp4pb@remail.nih.gov				Qy	672	ATGAAAAGACTGGSTATGTTCAATTACATGCCAAAAAATAAGTCTGGTCTA						
Tissue Procurement: DCTD/DTP				Db	965	ATGAAAGCTGTTAGGGATTTACATCAGCAGCTGAAAGCTGTC						
CDNA Library Preparation: Life Technologies, Inc.				Qy	792	CAACATTAGAGAATGAGTACTGATGAACTTACATGAACTTGGT						
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNL)				Db	1025	CTGATAAAATCCCTGGCAACGCTCCTGGCCATCTCTGGCC						
DNA Sequencing by: Baylor College of Medicine Human Genome				Qy	852	CGTCGTAGCCTCTGCACTTACATGGTACTACGACTTAAGTCAA						
Sequencing Center				Db	1084	CTGATGAAATTCATGGTAACTTACATGGTACTACGACTTAAGTCAA						
Center code: BCM-HGSC				Qy	911	ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG						
Web site: http://www.hgsc.bcm.tmc.edu/cDNA/				Db	1145	AGAGGTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG						
Contact: villalon@open.tmc.edu				Qy	965	ATGAAAGCTGTTAGGGATTTACATCAGCAGCTGAAAGCTGTC						
Villalon, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,				Db	1185	CAACATTAGGTTTACATGGTAACTGGTAACTGGTAACTGGT						
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,				Qy	972	CAACATTAGGTTTACATGGTAACTGGTAACTGGTAACTGGT						
Muzny, D.M., Gibbs, R.A.				Db	1204	CAACATTAGGTTTACATGGTAACTGGTAACTGGTAACTGGT						
Clone distribution: MGC clone distribution information can be found				Qy	979	CAACATTAGGTTTACATGGTAACTGGTAACTGGTAACTGGT						
through the I.M.A.G.E. Consortium/LILNL at: http://image.liln.gov				Db	1205	ACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG						
Series: IRAK Plate: 21; Row: d Column: 6				Qy	985	AGGAGACCCGGGAGGAGGACTCCACGTTGAGGTGAACTGGGAAACTA						
This clone was selected for full length sequencing because it				Db	1206	CTGATGAAATTCATGGTAACTGGTAACTGGTAACTGGTAACTGGT						
passed the following selection criteria: matched mRNA gi: 6855600.				Qy	992	CAACATTAGGTTTACATGGTAACTGGTAACTGGTAACTGGT						
Location/Qualifiers				Db	1207	CAACATTAGGTTTACATGGTAACTGGTAACTGGTAACTGGT						
1. .1584				Qy	1004	CAACATTAGGTTTACATGGTAACTGGTAACTGGTAACTGGT						
source				Qy	1012	CAACATTAGGTTTACATGGTAACTGGTAACTGGTAACTGGT						
/organism="Homo sapiens"				Qy	1019	CAACATTAGGTTTACATGGTAACTGGTAACTGGTAACTGGT						
/db_xref="GI:15080499"				Qy	1026	CAACATTAGGTTTACATGGTAACTGGTAACTGGTAACTGGT						
/clone="MGC-9222 IMAGE:3859644"				Qy	1033	CAACATTAGGTTTACATGGTAACTGGTAACTGGTAACTGGT						
/tissue_type="Ovary, adenocarcinoma"				Qy	1040	CAACATTAGGTTTACATGGTAACTGGTAACTGGTAACTGGT						
/clone_id="MGC_66"				Qy	1047	CAACATTAGGTTTACATGGTAACTGGTAACTGGTAACTGGT						
/lab_host="DH10B"				Qy	1054	CAACATTAGGTTTACATGGTAACTGGTAACTGGTAACTGGT						
/note="vector: pCMV-SPORT 6"				Qy	1061	CAACATTAGGTTTACATGGTAACTGGTAACTGGTAACTGGT						
/codon_start=1				Qy	1068	CAACATTAGGTTTACATGGTAACTGGTAACTGGTAACTGGT						
/product="Similar to serine (or cysteine) proteinase				Qy	1075	CAACATTAGGTTTACATGGTAACTGGTAACTGGTAACTGGT						
inhibitor, clade A (alpha-1 antiproteinase, antitrypsin),				Qy	1082	CAACATTAGGTTTACATGGTAACTGGTAACTGGTAACTGGT						
member 1"				Qy	1089	CAACATTAGGTTTACATGGTAACTGGTAACTGGTAACTGGT						
/protein_id="PAAH1191.1"				Qy	1096	CAACATTAGGTTTACATGGTAACTGGTAACTGGTAACTGGT						
/db_xref="GI:15080499"				Qy	1103	CAACATTAGGTTTACATGGTAACTGGTAACTGGTAACTGGT						
/translation="MPSSVWSWGTLLAGLCLVPSLAEDPOGDAAQKTDTSHHDDH				Qy	1110	CAACATTAGGTTTACATGGTAACTGGTAACTGGTAACTGGT						
PTENKTPMLAEPAFLSLRQLAHSNTNIFPSVSTATAAMISGTKDTHDELE				Qy	1117	CAACATTAGGTTTACATGGTAACTGGTAACTGGTAACTGGT						
GLAENLTPLEAQHESFQLRINTNQDQLQLTGNGFLUSEGLKLFLEDKK				Qy	1124	CAACATTAGGTTTACATGGTAACTGGTAACTGGTAACTGGT						
LYSEAFKWWNEDTEAKKQINDYVERGTOGTIVDLYKELDRTVPAVLYK				Qy	1131	CAACATTAGGTTTACATGGTAACTGGTAACTGGTAACTGGT						
ERPEPKDDEEDHFDVQDWTIVKVPMLNQHCKLSSNWLKMLGNRATAIF				Qy	1138	CAACATTAGGTTTACATGGTAACTGGTAACTGGTAACTGGT						
ERPEPKDQHLENELBHDITKFLENDRSALSNLHPLKSLTGTYDJKWSQLGK				Qy	1145	CAACATTAGGTTTACATGGTAACTGGTAACTGGTAACTGGT						
VFSNGADISGVTEAPKLSKAVKAVLTDKGTEAGAMFLEAPMSIPEVKFNK				Qy	1152	CAACATTAGGTTTACATGGTAACTGGTAACTGGTAACTGGT						
PFVFLMDONTKSPFLMGKVNPTQ"				Qy	1159	CAACATTAGGTTTACATGGTAACTGGTAACTGGTAACTGGT						

Db	1265	GTCACAGAGGAGGGACCCCTGAAAGCTCICCAAGGCCGATAAGGCTGTGCTGACCATC	1324	Db	481	GTTGAGAAGGGTACTCAAGGAAAATTGTGGATTGTGGCTAAGGACCTGTACAGAGACACA	540	
Qy	1032	GATGAAAGGGTACCGGGCCGGGCGCTATGTTCTGGAGCTATTCAATGAGCATT	1091	Qy	552	GTCTTCGCACTGTTACTATTTTCAGGTTAAGTGGGAACTGCTCTTGAGGTT	611	
Db	1325	GACGAAAGGGACTGAGCTGTGGGGCCATGTTAGGGCCATAACCCATGTCATC	1384	Db	541	GTTTTGCTCAGTGTGATTACATCCTCTTAAAGCCAATGGGAGACCCCTTGAGATC	600	
Qy	1092	CCACCAAGGAGTAAATTATAAACCATTCGTTCTGATGATCGAGGAACACTAA	1151	Qy	612	AAAGATACTGAGGAGAGATTCTAGTGTACAGTTACTGTCAAGTCCAAATGCAATG	671	
Db	1385	CCCCCGGGTCAAGTCAACAAACCTTGTCTTATGATGACAAATACCAAG	1444	Db	601	AAGGACACCGGAAAGGACTTCCACGTGACCGTGACCCAGTGGGCPATG	660	
Qy	1152	AGCCCATTTGTTAGGGTCAACCAACTCAGAA	1192	Qy	672	ATGAAAGACGCTGGTATGTTCAATATCACAATTGCAAAATAATGATGCTTAA	731	
Db	1445	TCTCCCCCTTCATGGGAAAGGGTGAATCCACCAAA	1485	Db	661	ATGAGCGTTAGGATGTTAACCCAGACTGTAAAGCTGTCAGCTGGTGTCTG	720	
REFERENCE	7			Qy	732	TAAATGAGTTTGTGTAACGCTATTGCTACTGCTATTGTTTACAGAGGTTAGCTT	791	
AR111410	AR111410	1185 bp	DNA	Qy	721	CTGATGAAATACCTGGCAATGCCACGCCATCTCTGTAGGGGAAACTA	780	
LOCUS	AR111410	Sequence 2 from patent US 6127145.	linear	Qy	792	CAACATTAGAAATGAGTTGACTCATGACATTACTAAATTGAGAACGAGG	851	
DEFINITION	AR111410			Db	781	CAGCACCTGGAAAAATGAACTCACCAGTATCATCACAAGTCCGGAAATGAGAC	840	
VERSION	AR111410.1	GI:12828258		Qy	852	CGTCGAGCGTTCTGACCTGCAAGTTAGTACCCGGTACTTACGACTTAAA	911	
SOURCE	AR111410	Unknown.		Db	841	AGAAGCTCTGCCAGCTACATTACCCARACTGTCCATACTGAACCTATGTCGAAG	900	
ORGANISM	AR111410	Unclassified.		Qy	912	TCTGTTAGGCACTAGTACCAAGTTACGTTAGTACCCGGTACTTACGACTTAAA	971	
REFERENCE	1	(bases 1 to 1185)		Db	901	AGCGTCTGGTCACGGCATCATAGGTTCACAAATGGGTCACCTCGGG	960	
AUTHORS	Sutliff, T.D. and Rodriguez, R.L.			Qy	972	GTTACTGAAGGCTCATTAAATTGAGTAAAGCTGTTCAAAAGCGCTTAACTATT	1031	
TITLE	Production of alpha ¹ -antitrypsin in plants			Db	961	GTCAAGAGGGACCCCTGAAAGCTCICCAAGGCCGCTAAGGCTGACCATC	1020	
JOURNAL	US 6127145-A 2 03-OCT-2000;			Qy	1032	GATGAAAGGGTACCGGGCCGCTATGTTCTGGAGGCTATTCCATGAGCATT	1091	
FEATURES	Location/Qualifiers			Db	1021	GACGAAAGGGACTGACCTGCTGGCCCATGTTAGGCCATATGCTATC	1080	
source	1..1185	/organism="unknown"		Qy	1092	CCACCCAGACTTAATTAACCATCGTTCTGATGATGAGCAACACTAA	1151	
BASE COUNT	328	a	324	Db	1081	CCCCCGAGGTCAAGTCAACAAACCCCTTGCTCTTAATGATGCAACAAATCCAA	1140	
ORIGIN	328	c	283	Qy	1152	AGCCCATTTATGGTAAGGTGTCAACCAACTCGAGAA	1192	
Qy	132	TTAGCTCATCAAAGTAATTCTACTAACATTTTTGTCTTATGCCACTGT	191	Db	1141	TCTCCCTCTTCAGGGAAATGCTGTTACCCCAAA	1181	
Best Local Similarity	28.1%	Score 429;	DB 6;	RESULT	8			
Matches	711;	Conservative 0;	Mismatches 470;	LOCUS	AX335339	1,352 bp		
Qy	12	GANGACCCTAAGGGACCGCGCCTCAAAACCGACCAAGCTATCGACCAAGACCAT	71	DEFINITION	Sequence 5848 from Patent WO1994629.	DNA	linear	
Db	1	GAGGATCCCGAGGAGATGTCGGGAGATGCCACCATGATACCCACATGAGTCAC	60	ACCESSION	AX335339		PAT 09-JAN-2002	
Qy	72	CGACATTAAATTACTCAAATTACTCCAAATTAGCGGAATTGCTTTCTTGATAGACAA	131	VERSION	AX335339.1			
Db	61	CCAACTTCAACGACGACGACGACGACGACGACGACGACGACGACGACGACG	120	KEYWORDS	GI:18126058			
Qy	132	TTAGCTCATCAAAGTAATTCTACTAACATTTTTGTCTTATGCCACTGT	191	ORGANISM	human.			
Db	121	CTGGACACAGTCACAGCAACCAATTCTTCCTCCAGTACGCTACAGCC	180	ORIGIN	Homo sapiens			
Qy	192	TTGGCATGTTAGTACTAACCGCATACCGATACCGATAGTAAAGTTTA	251	REFERENCE	Bukaoza; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Db	181	TTGGAATGCTCTGGGAGCAAGCTGACTGACGTGAAATCCGGGGCCTG	240	AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
Qy	252	AACTTAAATTGACCGAAATCCAGAAATCCAGAAATTCAGAGGTTTCAGAGTGTG	311	TITLE	Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,			
Db	241	AATTCAACCTCAGGAGATCCGGAGATCCGGGCTCGATCCATGAGGTCTCC	300	KEYWORDS	Horigan, S., Scoppe, D.R. and Weaver, Z.			
Qy	312	AGACTCTGAATCAACCTGATCTCAATTGATTAATCTACTGTGGTTATTG	371	JOURNAL	Cancer gene determination and therapeutic screening using signature			
Db	301	CGTACCCCTAACAGCCAGACGCCAGACGCCAGCTCCGCTACACCGGAAATGGCTCC	360	FEATURES	Patent: WO1994629-A 5848 13-DEC-2001;			
Qy	372	TCTGAAGSTTTAAATTGGTGTGACAAATTCTTAGAGACCTCAAAATCATGT	431	LOCATION/QUALIFIERS	Avalon Pharmaceuticals (US)			
Db	361	AGGGAGGCCTGAGCTAGTGTGAAATGGGATTTGGAGGATTTGACCTCA	420	ORIGIN	1..1352			
Qy	432	GAGGCTTCACTGTCACCTCGGGACACCAAGGCCAGAAAGATGATAC	491	ORGANISM	/organism="Homo sapiens"			
Db	421	GAAGGCTTCACTGTCACCTCGGGACACCAAGGCCAGAAAGATGATAC	480	DB_XREF	/db_xref="taxon:9606"			
Qy	492	GTGAGAAGGGACCCAGGTAAGATCGTGAACCTAGTGTGATACC	551	BASE COUNT	349 a	386 c	325 g	
				Query Match	28.1%	Score 429;	DB 6;	Length 1352;

Best Local Similarity 60.2%; Pred. No. 2.6e-81; Matches 711; Conservative 0; Mismatches 470; Indels 0; Gaps 0;	Db	11112 GACCGAGAAAGGGACTGAAGCTGGGGCATGGCCATACCCATGTCATC 1171
Qy 12 GAGACCCCTCAGGGGACGCGCTCAAAACGCCAACCCAGTCATCAGCACAGACCAT 71	Qy	1092 CCACCAAGAGTTAAATTAAACCATTCGTTTTCGATGATCCAGAGAACATAAA 1151
Db 92 GAGGATCCCGGGAGATGCTGCCAGAAAGACATCCACATGAGATCAC 151	Db	1172 CCCCGGCGGTCAAGTOAACAAACCTTTCGTCCTTAATGATTGAAACAAATACCAAG 1231
Qy 72 CGACATTATAAAATTACTCCAAATTINGCCGATTTGCTTCTGTAGAGCAA 131	Qy	1152 AGGCCATTTATGGTTAAGGTGCTCAACCCAACTCAGAA 1192
Db 152 CCAACCTTCACAAAGTCACCCCAACCTCGTGTAGTCCCTCAGCTTACGCCAG 211	Db	1232 TCCUCCCTCTTCTATGGAAAAGTGATCCACCAAA 1272
Qy 132 "TAGCTGATCAAAGATTCATCAACTAACATTTTAGTCCTGTTCTATTGCCACTGCT 191	RESULT 9	
Db 212 CTGGCACACAGTCCACAGAACATCCAAATATCUTCTCTCCAGTCAGCATGCC 271	HUMA1ATM	
Qy 192 TTGCCATGTTAGTTAGGTTACTAAAGCCGATACCATGAGATTAGAGGTTA 251	LOCUS	
Db 272 TTGCAATGCTTCCCTGGGACCGCTGACACTCACGTAAATCCCTGGAGGCCCTG 331	DEFINITION	
Qy 252 AACTTAAATTGACCGAAATCCTCAGAAGGGTTTCAAGAGTTGTTG 311	Human alpha-1-antitrypsin mRNA, complete cds.	
Db 332 AATTTCACCTTCAGGATTCGGGATCTAGATCATGAGGTTCAGAGCTTCTC 391	ACCESSION	
Qy 372 TCGAAGGTTAAATTTGTTGACAATTCTCAATTGCAATTAAACTACTGGTAACCGGTTTATTG 371	K01396_1	
Db 452 AGGAGGGCCCTGAAACGGCACAGCCACCTCAGCTGACCCATGGCTGTCCTC 511	VERSION	
Qy 432 GAGGCTTTACCGTTAATTTGGTACTGGAACTGAAATTAAATGATTGATT 491	K01396_1	
Db 512 GAACCTCTACGTGCAATTGGGAAACCCGAGGCCAAGAAACAGATAACGATTAC 551	KEYWORDS	
Qy 492 GTGAGAAAGGCCACCGGTAAGATCCTGGACCTAGTTGGGATGTTAAAGGACACA 631	alpha-1-antitrypsin; antiprotease; antitrypsin.	
Db 572 GTGAGAAAGGTACTAACGGTAAATGTTGGTCAAGGAGCTTGAGAGACACA 631	SOURCE	
Qy 552 GTCTTGCACTAGTTAACTATTTTCTCAAGGGTAAGTGGAAACCTCTCTGGAGGTT 611	ORGANISM	
Db 632 GTTATTGCTCTGGGAATTAACATCTCTTAAAGGAAATGGAGAACCCTTGGAGTC 691	Homo sapiens	
Qy 612 AAAGATACTGAGGAGAACATTTCAGTCAAGTTACTCATGTCAAATGTCAAATG 671	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Db 692 AAGGACACGGAGGAAGGAACTCCACCTGGACCGGTGAAATTGAGT 751	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
Qy 672 ATGAAAAGACTGGGTATGTTCAATTCACATTGCAAAATTAAAGTTCTGGGCTTA 731	DEFINITION	
Db 752 ATGAGCGTTTAGGCATGTTACATCCAGCACTGTAAGGAGCTTGAGTC 811	Human alpha-1-antitrypsin mRNA.	
Qy 732 TTAAAGAAGTTAGGTAAGCTGACTCATGACATTACTAAATTTAGAACAGGAGT 791	ACCESSION	
Db 812 CTGAGAAATACCGGCAATGACCTCTCCGCACTTGCTGTAGGGGAACAA 871	K01396_1	
Qy 792 CAACATTAGAGAATGAGTTAGCTGACTCATGACATTACTAAATTTAGAACAGGAGT 851	VERSION	
Db 872 CAGCACCTGGAAAATGAACCTACCCACGATATCACCAAGTGTGAAATTGAGC 931	K01396_1	
Qy 852 CGTCGTAGGGCTCTGGCTCAACTGGCATCTAGTCCTGGGCGAACCTCTCCGG 1051	KEYWORDS	
Db 932 AGAAGGTCCTCCAGCTTACATTACCCAAACTGTCCCATATCTGAACTGAG 991	alpha-1-antitrypsin.	
Qy 912 TCTGTTTAAAGCCAGTTAGGTTACCAAGTTTCTACGCTGGCCATTGAGGTT 971	SOURCE	
Db 992 AGCGTCTGGCTCAACTGGCATCTAGTCCTGGGCGAACCTCTCCGG 1051	Organism="Homo sapiens"	
Qy 972 GTTACTGAAAGAGCTCATAAAATTGAGTAAGGTGTCACAAAGCCCTTAACATT 1031	/codon_start=1	
Db 1052 GTCACAGGAGGCACTCCGCTGAAGTCTCCAGGCGTCAATANGGCTGACCATC 1111	/protein_id="AAE59375_1"	
Qy 1032 GATGAAAAGGTACCGAGGCCGCCGCTGCTGAGCTTCCAAATGAGCAT 1091	<1..1352	

ERPFEYKDEEEEDFHQVTTVKVPMKMRKGMENTIOHCKKLSSNVLLMKYLNATAIF
FLPDEGFLQHIEENEIETHDITIKFELNEDRISASLHLPKLSITGTYDLKSVLQLGQITK
VFSNGADLSGCTTEAFLKLSAVHKAFLTIDEKGTEAAGMFLAEIPMSIPPEVKENK
PFVFLMIEQNTKSPLFMKVYNNPTQR" 20 .91

sig_peptide
/gene="PI"
/note="alpha-1-antitrypsin signal peptide"
92 .1273
mat_peptide
/gene="PI"
/product="alpha-1-antitrypsin mature peptide"
241
old_sequence
/gene="PI"
/citation=[1]
286
conflict
/gene="PI"
/citation=[3]
/replace="n"
old_sequence
435 .439
/gene="PI"
/citation=[1]
748
old_sequence
/gene="PI"
/citation=[1]
784
old_sequence
/gene="PI"
/citation=[1]
787
old_sequence
/gene="PI"
/citation=[1]
799 .802
old_sequence
/gene="PI"
/citation=[1]
837
old_sequence
/gene="PI"
/citation=[1]
1114
old_sequence
349 a
386 c
90 bp upstream of DdeI site.
BASE COUNT 325 g 292 t
ORIGIN 28.1% Score: 429; DB: 9; Length: 1352;
Best Local Similarity 60.2%; Pred. No. 2.6e-81;
Matches 711; Conservative 0; Mismatches 470; Indels 0; Gaps 0;

Query Match
12 GAGAGCCCTTAGGGGACGCCGCTAAAAACCCACCCAGCTTACCCAGAGCCAT 71
Db 92 GAGGATCCCAGGGGAGATGGTCCCCAGAAGACAGATACTCCACCATGAGATCAC 151
Qy 72 CGACGTTTAAATAAAATTACTCCAAATTAGGCCGAATTGCTTTCTGTATAGACAA 131
Db 152 CCAACCTTCAACAGCTTACCCCAACCCGCTTACCCGCAAGCTTACCCGAG 211
Qy 132 TTAGCTCATAAAGTAATTCTACTAACATTTCCTGTTTCTGTTCTPATGCCACTCT 191
Db 212 CTGGACACAGTCACAGCACAATTCCTCCTCCACAGGATCGTACAGCC 271
Qy 192 TTGCGCATGTTGATTTAGCTAAAGCCGATACCCATGAGGAAATTAGGGTTA 251
Db 272 TTGCAATGCTCTCCCTGGGACCAAGGTCACACTCAGATAAAATCTGGGGCTG 331
Qy 252 AACTTTAATTGACCCGAATCCAGAAATTAGGCCGAATTGCTTTCTGTATAGACAA 131
Db 332 AATTCAACCTCAGGAGATCCGGAGTCACATGAGGTTTCAGAGTTCTG 311
Qy 312 AGAACTTGAATCAACCTGCAATTGCAATTAACTACTCTGAAAGCTTCAGGAACCTC 391
Db 392 CGTACCCCTAACCCGCAAGCCAGGCACTGACCCGCAATGGCTGTCCTC 451
Qy 372 TCTGAAGTAAATTGTTGACCCGAATTCCTGAAATTCTGAGACTATCATAGT 431
Db 452 AGGGGGCTGAGCTACTGGATAAGTTGGAGATGTTAAAGGTGACCACTCA 511

Qy 432 GAGGCTTTACCGTTAATTGTTGTTGACTCTGAGGAAGCTAAAGCAAAATTAAATGATTAT 491
Db 512 GAAGCCCTCAGTCGTAACCTCGGGGACACCGAACGCCAGAATCACAGGATAC 571
Qy 492 GTTGAGAAAGGCCAGGTTAAGATGTTGACTCTGAAATTAGATCGTGATACC 551
Db 572 GTGGAGAAAGGTACTCAAGGAAATTGTTGAGCTTGACAGAACACA 631
Qy 552 GTCCTCGCACTAGTTAATCTACATATTTCAGGAAACGNCCTTTCGAAAGTT 611
Db 632 GTTTTGCTCTGGTAATTACATCTTAAAGCAATGGGAGAACCCMTGAGTC 691
Qy 612 AAAGATACCTGAAAGGAAGATTTCATGTTGACAAAGTTACTCTGCAAGTTCCATG 671
Db 692 AAGGACACCGAGGAGGAGACTTCACCGTGGACCCAGTGACCGTGAAGGTGCCCATG 751
Qy 672 ATGAAAAGACTGGGTATGTTCAATTATCAACATTGAAAAAATTAAAGTCTGTTGGCTTA 731
Db 752 ATGAGCGTTAGSCATTTAACATCCAGCAGTAACTGAAAGCTGTCAGCTGGTCTG 811
Qy 732 TTATGAACTTGGAAAGTGTGACTCATGACATTAACTAAATTTCAGAACGAGAT 851
Db 812 CTGTGAAATACCTGGCAATGCCACGCCATCTTCCTGCTGTAGGGAAACTA 871
Qy 792 CAACATTGAGAATGGTGTGACTCATGACATTAACTAAATTTCAGAACGAGAT 851
Db 872 CAGACCTGAAATGAACTCAACCCACATATCACCAACTTCCAGAAATGAGAC 931
Qy 852 CGTCGTGGCCTCTCTGACCTGCAAGGTTAAGTTCACCGGTATTTAGGACTTAAA 911
Db 932 AGAAGGTCTGGCAAGCTTACCAAACTGTCCTTACTGAAACCTATGTCAG 991
Qy 912 TCGTTTTRAGGCCAGTTAGTTACCAAGTTTCTAAAGGTTCTAACGGGGAFTGAGTTG 971
Db 992 AGCTCTGGTCAACCGGCATCACTAAAGGCTTCTGACAACTGGCTCTCGGG 1051
Qy 972 GTTACTGAGAAACTCCATTAAATGGTAAAGCTTCAAAAGCCGTTCTTAACATT 1031
Db 1052 GTCACAGGGGACCCCTGAGCTTCCAGGGCTTCAAGGGCTTCTGACCT 1111
Qy 1032 GTGAAAAGGTCTCGAGGCCGGCTATGTTCTGGACTATCCATTGAGCAT 1091
Db 1112 GAGGAGAAGGGACTGAGCTGCTGGCCATGTTTAGGCCATACCCATGTTAC 1171
Qy 1092 CCACCAAGGTTAAATTAAATAACCATTCCTTTCTGATGCGAAACACTAA 1151
Db 1172 CCCCGGAGGCAAGTCAACAAACCCCTTGTCCTCTTAATGATTGACAAAATCCAG 1231
Qy 1152 AGCCCATGTTATGGTAAGGTTGCAACCCAACTCAGAA 1192
Db 1232 TCTCCCTCTTCTGAAAAGTGTAAATCCACCCAAA 1272

RESULT 10
AX335338
LOCUS Sequence 5847 from Patent WO1994629.
DEFINITION 1371 bp
ACCESSION AX335338
VERSION AX335338.1
KEYWORDS GI:18126057
SOURCE human
ORGANISM Homo sapiens
SOURCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horigan,S., Soppet,D.R. and Weaver,Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
PATENT WO 0194629 A 5847 13-DEC-2001;
JOURNAL Avalon Pharmaceuticals (US)
FEATURES Location/Qualifiers
Source 1 . 1371

BASE COUNT	ORIGIN	Query Match	Score	DB	Length
350 a		28.1%	429;	6;	1371;
Best Local Similarity		60.2%	Pred. No. 2.6e-81;		
Matches 711;	Conservative	0;	Mismatches	470;	Indels 0;
					Gaps
Qy	12	GAGACCTCAAGGAGGCCCTCAAAAACCGACACCAGTCATCACGACCAAGACCAT	71		
Db	111	GAGTCTCCCGGGAGATGCTGCCGAGACATACATGATCAGATCAC	170		
Qy	72	CCGACTTTAATAAATTACTCCAAATTAGCCGAATTGGCTTCTGTATAAGACAA	131		
Db	171	CCACCTCAACAGATACCCCCAACCTGGCTGACTTCGCCTTAACGCCAG	230		
Qy	132	TTCAGTCATCAAGTAAATCTACTACATTTTTAGTCCTGTTCTATTGCCACTGCT	191		
Db	231	CTGCAACACCAGTCCAAACAGACCAATTATCPTCTTCCCAGTAGGCATACAGCC	290		
Qy	192	TTCGCCATGTTGAGTTAGGTACTAAAGCGATAACCATGCGAAATTAGAGGTTTA	251		
Db	291	TTTCGCAATGCTTCCCTGGGGACAGGCTGACACTACAGATGAAATCCUGGGCCCTG	350		
Qy	252	AACCTTAATTGACCAAATCCAGAACGCCAAATTCAAGGAGTTTCAGAGTTGTTG	311		
Db	351	AATTCAACCTCAGGAGATTCCGAGATTCAGATCATGAGGCTTCAGGAACCTCC	410		
Qy	312	AGACATTGAAATCACCTGATCTCAATTGCAATTAACTACTGTAACGTTTATTTTG	371		
Db	411	CGTACCCCAACCCAGGAGACACCCAGTCCAGCTGACCGGAATGGCTCC	470		
Qy	372	TCTGAGGTTAAATTGTTGCAATTCTAGAACGCTCAAGAACATATCATAGT	431		
Db	471	AGCGGGCTGAGCTAGTGTGATAATTGGAGATGTTAAAGAGTTACCACTCA	530		
Qy	432	GAGGTTTACCGTTAATTGGTGTAACTGAGGAACCTAAAGGAATAATTGTTAT	491		
Db	531	GAAGCTTCAACTGTCAACTTGGGAAAGGACTTCACTGGGACCCAGAAACAGATAACCGTTAC	590		
Qy	492	GTGAGAAAGGCCACCGGGTAAGATCGTTGACTCTAAAGAAATTAGATCGTGTAC	551		
Db	591	GTGGAAAGGGTACCAAGGAAATTGGGATTGGCTCAAGGAGCTTGTGAGCACCA	650		
Qy	552	GTCTGCACTAGTTAACTATATTTCAGGGTAAGTGGGAACCTTTCAGGTT	611		
Db	651	GTTTGTGCTGGAAATACATCTTAAAGGAAATGGGACCCUTTGAGTC	710		
Qy	612	AAAGTACTGAAAGGAAGATTTCATGTTGATCAAGTTACTACTGTCAAAGTCATG	671		
Db	711	AGGGCACCGGGAAAGGACTTCACTGGGACCCAGGTGACCCGTAAGGTGCTTATG	770		
Qy	672	ATGAAAGACTGGTATGTTCAATTACATGAAATTAACTGCAAATTAATTGTT	731		
Db	771	ATGAGCTGGTTAGGGATGTTAACCTCAGCACTGTAAGACTGCTTGTGAGCT	830		
Qy	732	TAAATGAAAGTATTAGTAAACGCTTAATCCAGCACTGTAAGACTGCTTGTGAGCTT	791		
Db	831	CTGATGAAATACCGGGCAATCCGGCACTTCTCGTGTGAGGGAAACTA	890		
Qy	792	CAACATTAGAGAAATGAGTGTGACTCATGACATTATACCTAAATTGAGAACGGGAT	851		
Db	891	CAGCACCTGGAAAATGAACCTACCCACGATATCACACCAAGTCTGGAAATGTGAGAC	950		
Qy	852	CGTCGATGCCCTCTGACCTGCCAGTTACATTACCCAAACTGTCATGAGCTTACCACTAA	911		
Db	951	AGAAAGTCTGGCTCAGTTACATTACCCAAACTGTCATGAGCTTACCACTAAATGTGAGAC	1010		
Qy	912	TCTGTTTACGCCAGTTAGTAACTACCGGCACTTCTCGTGTGAGGGCTG	971		
Db	1011	AGCGTCTGGTCAACTGGCATCTAGCTTCACTGGCTGACCTCTCGGG	1070		

LYHSEAFPTVNGFDTTEAKQINDYVEKKGTOKGKIVDLYKELDRDTVFAVNYYIFFKGKW EPRFEVDETEEDFHQDQVTVKVKPMKMRLEMENIOHKCKLSSWVILMKYLGNATAIF FLPDEQFLQHLENELDITKFLNEDRSASLHPKLSITGTDLKVSLGQLCITK VFSNGADLSGIVTEAIPKLSRAVKHAVLTIDEKGTEAAGAMFLEAPMSSPPEYKFEN PFPVFLMIBQNTKSPLEMKGKVNPQT	Db	951	AGAAGGCTCTGCAGCTTACATTACCAACTGTCATTAACCTGATCTGAG	1010
sig_peptide	Qy	912	TCTGTTTAGGCCAGTAGTTACCAAGTTTCTPACCGSTGCCATTAGTCTGAG	971
mat_peptide	Db	1011	AGCGRCTGCGTCACATGGCATCATAGGTCTCAGAAATGGCTGACCTCTCCGG	1070
POLY_A_site	Qy	972	GTACTGAAGAGCTCATAAATTGAGTAAGCTTCACAAAGCGCTAACATT	1031
BASE COUNT	Db	1071	GTCACAGAGGGCACCCCTGAAGCTCCAGGCGTCATARGCTGACCAT	1130
1371	Qy	1032	GATGAAAGGGTACCGAGGCCGCGCTATGTCCTGAAAGCTTCAATGAGCATT	1091
350 a	Db	1131	GACGAGAAAGGACTGAAGTGTGGCCATACGGCTACCCAGTCAT	1190
388 c	Qy	1092	CCACCGAAGTTAAATTAAACCATTCGTTCTGATGATGGAGAGAACCTAA	1151
340 g	Db	1191	CCCCCGAGGTCAAGTCACAAACCTTGTGCTCTTAATGATGAAACAAATACCAAG	1250
293 t	Qy	1152	AGCCATTGTTAGGTAAAGGTAAAGTGTCAACCCAACTCAGAA	1192
	Db	1251	TCTCCCTCTCATGGAAAAGTGTGATCCACCAAA	1291
Query Match	Score	28.1%	Score 429; DB 9; Length 1371;	RESULT 12
Best Local Similarity	60.2%	DB 2, 6e-81;	DB 192	AK026174
Matches	711;	Indels 0;	Qy 192	LOCUS
Conservative	0;	Gaps 0;	CCGCCATGTTAGCTTACGGCTCAAAACGCCACCAAGTCACGCCAAAGCATT	Hom sapiens cDNA: FLJ22521
Matches 711:	0;	Pred. No. 2, 6e-81;	251	bp mRNA
		Missmatches 470;	TTGCAATGCTCTCCCTGGGCCAACGGCTGACACTCACCGATGAAATCTGGAGGCCCTG	Clone HRC12316, highly similar to
		Indels 0;	350	AF113676 Homo sapiens clone FLB2803 PRO0684 mRNA.
		Gaps 0;	Qy 291	ACCESSION
			AK026174	VERSTON
			AK026174.1 GR:10438937	KEYWORDS
			oligo capping; fls (full insert sequence);	SOURCE
			Hom sapiens primary human renal epithelial cells cDNA to mRNA,	ORGANISM
			clone.lib:HRC clone:HRC12386.	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
			Qy 252	REFERENCE
			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	AUTHORS
			Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Okamoto,S., Okitani,R., Ota,T., Suzuki,Y., Obayashi,M., Nishi,T.,
			Db 351	JOURNAL
			AAATTCAACCTCACGGGATTCGGGGCTAGATCCATGAGGCTCCAGAACACTCCCTG	Unpublished
			410	1 (sites)
			Qy 312	REFERENCE
			AGACCTTGATCAACCTGATCTCAATTGCAATTAACTCTGGTAACGGTTTATTTG	AUTHORS
			371	1 (sites)
			Db 411	JOURNAL
			CGTACCCCTAACCGCCAGCAGCCAGCTAACCGCCAGTACGGCTCC	1 (sites)
			470	REFERENCE
			TCTGAAAGTTAAATTTGGTACACATTCTAGAAGACGTCAGAAACTATATCATGT	AUTHORS
			431	1 (sites)
			Db 471	JOURNAL
			ACGAGGGCTGCAAGCTAGTGGATAAGTTGGAGGATSTAAAAGTTGTAACACTCA	1 (sites)
			530	REFERENCE
			Qy 432	AUTHORS
			GAGGCTTTACCGTTATTGGTACAGTGAGGCTAAAAAGCAATTAAATGATTAT	1 (sites)
			491	JOURNAL
			Db 531	Shibahara,T., Tanaka,T., and Nakamura,Y.
			GRAAGCAGCCACCGGACAGCAGCTAACCTGGGACACGGCAAGACAGTCAGATTAC	1 (sites)
			590	COMMENT
			Qy 492	Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center, Shirokanedai, 4-6-1, Minato-ku, Tokyo 108-0039, Japan (E-mail: chal@ims.u-tokyo.ac.jp, Tel: 03-5449-5286, Fax: 03-5449-5116)
			Db 591	NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5' - 6' - 3' - end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).
			Qy 552	FEATURES
			GTCCTGGCCTACAGTAACTATATTTCAGGGTAAGTGGAACTGCTCTGGTCTTA	Source
			651	1. 1399
			GTGAGAAGGGTACTAACGGGAAATTGGTACAGTTGAGGCTTGTGAGGACACA	/organism="Homo sapiens"
			650	/db_xref="Taxon 9606"
			Qy 651	/clone="HRC12386"
			GTGTTGCTCGGTGATTACATCTCTTAAGGCAAATGGAGAACCTTGTGAAACTC	/cell_type="primary human renal epithelial cells"
			710	/clone_lib="HRC"
			Qy 612	/note="Cloning vector pME18SFL3"
			AAAGATCTGAGGAAAGATTTCATGTTGATCAAGTTACTCTGAAAGTCCAATG	1. 1199
			671	/note="highly similar to AF113676 Homo sapiens clone
			770	FLB2803 PRO0684 mRNA"
			772	BASE COUNT
			ATGAAAGCTGGTATGTCATATTACACATTACACATTACACCTGGCTCTTA	378 a
			731	329 g
			ATGAAAGCTTAGGGCATGTTAACATCCACACTGTAAGAGCTGGCTGGTCTG	297 t
			830	ORIGIN
			Qy 732	
			TTAATGAGTATTAGGTAACTGCTACTGCTATTTCAGGAAAGCTCCCTGGAGGT	
			791	
			Db 831	
			CTGAGAAATACCTGGCATGTCACCGCCATCTCTCTGCTGATAGGGAAACTA	
			890	
			Qy 792	
			CAACATTAGAGAATGAGTGTACTCATGACATTACTAAATTTAGAAACGAGGAT	
			851	
			Db 891	
			CAGCACCTGGAAAATGAACTCACCCACGATATCATCACCAAGTCTGGAAATGAAAC	
			950	
			Qy 852	
			CGTCGTAGGGCTCTCTGCGACCTGGCAAAGTTAAGTATCACGACTTAA	
			911	

Query Match	Score 249; DB 9; Length 1399;	Db	1136 GACGAAAGGACTGAAGGTGCGGGCATACCGATACCGATACCGTAC 1195
Best Local Similarity	60.28%; Pred. No. 2.6e-81;	Qy	1092 CCACGAAAGTTAAATTATAACCATGCGGATGCGGCAACTCAA 1151
Matches	711; Conservative 0; Mismatches 470; Indels 0; Gaps 0;	Db	1196 CCCCCGGTCAAAACCTTCAGGAAATAATGAAATACCG 1255
Y	12 GAAAGACCTCTAAGGCGAGCGCCCTCAAAAAACCGAACCCAGAACCCAT 71	Qy	1152 AGCCCATGGTTAGGTAAAGGTCAACCCAACTCGAA 1192
b	116 GAGGATCCCAAGGAGATGCTCCAGAAAGAATACCCATGATCAGGATCAC 175	Db	1256 TCTCCCTCTCATGGAAAATGGTGTGATCCATGCCAAAA 1296
Y	72 CGCACTTTAAATAATTACTCCAAATTAGCGGAATTGGCTTCTTGTATAGACAA 131		RESULT 13
y	73 CGAACCTTCACCAAGATCACCCCAACCTGGTGAAGTCCAGCTTACGCCAG 235	E00195	E00195 14 34 bp
b	132 TTAGCTCATCAAGTAATTCTACTAACATTTTAGTCCTGTTCTATGCCACTGCT 191	LOCUS	14 34 bp
b	236 CTGGCACACCGTCCACAGCCAAATATCCTTCCTCCCAAGTGCATCCAGGCC 295	DEFINITION	encoding human antitrypsin.
b	192 TTGCCCCATGTTAGTTAGTAACTAAAGCCGATACCAATGAGGATTAGGGTTTA 251	VERSION	E00195.1
b	296 TTGCTCAAGTCTCCCTGGGACAGGCTGACACTACGANGAAATTCTGGAGGCC 355	KEYWORDS	GI:2168491
b	252 ACTTTAATTGACCGAAATTCACGAGGGTTCAAGAGTTGTTG 311	SOURCE	JP 1984091886-A/1.
b	356 AATTCAACCTACAGGATTGGGAGCTCACATCCATGAAGGCTCCAGAACCTC 415	ORGANISM	Homo sapiens
b	312 AGAACTTGAATCACCTGATTCCTCAATTGCAATTAACTACTGTTAAGGGTTATTG 371	REFERENCE	(bases 1 to 1434)
b	416 CTGACCCCTAACCCAGCCAGACCCAGCTGACCTGACACCGCAATGGCTCCTC 475	AUTHORS	Guren, E. K. and Richiyoado, U.
b	372 TCTGAGGTTAAATTGTTGACAAATTCTTACAGACGTCAGAAACTATATCATAGT 431	TITLE	PROTEASE INHIBITOR FOR DEVELOPMENT OF
b	476 AGCGGGGCTGANGCTAGTGATAAGTGTAAAGCTGACCACTCA 535	JOURNAL	ADJUSTED PROTEIN
b	432 GAGGCTTTACCGTAAATTGGTGTATACTGAGGTAATAGCAATTATGATTAT 491	COMMENT	Patent: JP 1984091886-A 1 26-MAY-1984;
b	536 GAAGCTTCATGTCAACTCGGGACACCGGCAAGAAACAGATCACGATTAC 595	OS	JIMOSU CORP
b	492 GTTGAGAAAGCACCAGGGTAAAGATGTTGACCTGTTGATGACCC 551	PN	(human)
b	596 GTGGAGAAGGGTACTCAAGGGAAATTGTGGATTGGCTGAGGTGCTCATG 655	PD	PN 1984091886-A/1
b	552 GTCCTGGACTAGTAACTATTTTCAAGGGTAAGTGGAACGTCCTTGAGGT 611	PD	26-MAY-1984,
b	656 GTTTCCTGGTGTATTACATCCTTAAGGAAAGCTTGAAGTC 715	PF	JP 1983147907
b	612 AAAGATACTGAGGAGAATTCTGTTAATTCAGTTACTACTGTCAGGTTCCAATG 671	PR	28-APR-1983 US 83
b	716 AAGGACACCGGGAGAGGACTTCACGTCGACCGTACCGTGAAGGTGCTATG 775	PC	007H21/04, C12N1/16, PC
b	672 ATGAAAGACTGGTATGTTGAAATAATTAGTCTGGTCTTA 731	C12N15/00, A61K35/72, A61K37/64, C07G7/00, C12N1/16, PC	
b	776 ATGAAAGGTTAGGCTGTATGTTAACCTCAGGACTGAAAGCTGTCGGTGC 835	C12N9/99,	
b	732 TTAATGAGTATTAGGTAACCTGTTCAAGGGTAAGTGGAACGTCCTTGAGGT 791	PC	C12P21/00, (C12N1/16, C12R1:865);
b	836 CTTGATGAAATCCTGGCAATGGCAAGTCCAGGATATCAGAAATGGAAATGAGAC 895	CC	strandedness: Double;
b	792 CAACTTAGAATGAGTTGACTCATGACATTATTAATCTTGTATGGAAATGAGAC 955	CC	topology: Linear;
b	896 CACCACTGGAAAATGACTACCCAGATATCAGAAATGGTGTGAC 955	CC	hypothetical: No;
b	852 CGCGTAGCGCTCTGGCACCTGCAAGTAAAGTAAAGPATACCGGTACTAGCTAA 911	FT	sig_peptide 47..118
b	956 AGAAGGCTCTGCCAGCTTACGCCAAACTGTCCTAACTGGAATCTGATCAG 1015	FT	mat_peptide 119..1300
b	912 TCTGTTAGGCCAGTAAAGTAACTACCAAGTGGTTCTACGGTCCGATTTGAGTGGT 971	FT	/product="human antitrypsin".
Qy	Qy	FEATURES	Location/Qualifiers
Qy	Qy	SOURCE	1..1434
Qy	Qy	ORIGIN	/organism="Homo sapiens"
Qy	Qy	BASE COUNT	/db_xref="Taxon:9506"
Qy	Qy	Query Match	369 a
Qy	Qy	Best Local Similarity	342 c
Qy	Qy	Matches	340 g
Qy	Qy	Length	293 t
Qy	Qy	DB 6;	Score 2.6e-81;
Qy	Qy	Indels	0;
Qy	Qy	Gaps	0;

Qy	192	TTGCCATGGTGTAGTTAGCTAAACCGATAACCATGACGAGATTAGAAGGTTA	251	DEFINITION Sequence 3 from Patent EP 0137633.
Db	299	TTGCAATGCTCCCTGGACCAAGGTCACATCAGTAAATCCGGCCCTG	358	ACCESSION 104196 VERSTON 104196.1 KEYWORDS GI:591838 SOURCE Unknown.
Qy	252	AACTTTAATGACCGAAATCCAGAACGCCAAATTCACAGGGTTTCAGAGTTG	311	ORGANISM Unclassified.
Db	359	AACTTCACCTCACGGATTCGGGGCTAGATCAGTCATGAAAGCTCC	418	REFERENCE 1 (bases 1 to 1434) AUTHORS Parker, M.L. and Kawasaki, G.H. TITLE Method of expressing alpha-1-antitrypsin in bacteria and its use in therapeutic formulations, and vectors and bacteria for such method and their production PATENT: EP 0137633-A1 3 17-APR-1985; LOCATION/QUALIFIERS 1. ORGANISM="unknown" BASE COUNT 369 a 432. C 340 g 293 t
Qy	312	AGAACCTGAAATCAACCTGATTCATGGAAATTAACTACTGGTAACGGTTATTG	371	Query Match 28.1%; Score 429; DB 6; Length 1434; Best Local Similarity 60.2%; Pred. No. 2.6e-81; Matches 711; Conservative 0; Mismatches 470; Indels 0; Gaps 0;
Db	419	CCTACCCCTCAACCGCAAGCCAGCAGCCAGCTAACGGCAATGGCTGTC	478	Qy 12 GAAGACCTCAAGGGAGCCGGCTCAAAAACGCCAGTCAGCAAGAACCAT 71 Db 119 GAGGATCCCAGGGAGATGGCTGCCAGAGACAGATACCCACCATGATCAGGATCAC 178
Qy	372	TCTGAAGGTAAATGGTGTGACAAATTCTAGAGCCAAACTATATGATCTG	431	Qy 72 CCGACTTTAATAAAATTACTCCAAATTAGCGGAATTAGCGTCTTTGTATAGCAA 131 Db 179 CCAACCTAACAGATACCCCAACCTGGCTGAGTCGCTTACCCGCAACG 238
Db	479	ACGGAGGCCCTGAAGCTAGTGGATAGTTGGAGGTAAAGTGTACCACTCA	538	Qy 132 TTAGCTCATCAAAGTAATCTACTAACATTTTTGTCTATGCGTCT 191 Db 239 CTGGCAACCCAGTCACAGCCAAACAGCCAAATCTCCCTAGGACATCCCTACGCC 298
Qy	432	GGGCTTTACCGTTAAATTGGTGTATACTGGCAATGGAAATGGCTTGAAGTC	491	Qy 192 TTGCCCATGTTGAGTTGGTACTAAAGCCGATCCATGAGAGATTAGGTTA 251 Db 299 TTGCAATGCTCCTGGGCAACGGCTGACACTACGATGAATCTGGGGCTG 358
Db	539	GGAGCAGGAGGAAGGACTCCACGTGACCCAGGTGACCAACGGAC	598	Qy 252 AACUTTAATTGACCAAATCCGAAACGCCAAATTCACGAGGGTTCAAGAGTTGTT 311 Db 359 AATTCAACCTCAGGAGATTCGGGCTCAGATCAGGGCTCAGAACCTC 418
Qy	612	AAAGATACTGAGAGGAGATTTCTATGTTGATCAAGTTACTACTGTCAAAGTCCAAATG	671	Qy 312 AGAACTGAAATCAACCTGACATTAACTTAATGTTTACCAAGGAAAGGTAAAGCTT 371 Db 419 CGAACCTCAACAGCAGCCAGTCAGCTGCCAGTCAGCTGCTTCAGCTC 478
Db	779	ATGAAGCGTTAGGGATGTAACTCCACACSTAAGAAGCTGCCAGTGGCTG	838	Qy 372 TCTGAAGTTAAATGGTGCACAAATTCTAGAAAGCTCAAGAACATATATCATG 431 Db 539 GAGGCCCTCACGTCACACTCGGGACACCGAGGGCCAGAACGATCAACGATAC 598
Qy	732	TTAATGAAGTATTAGGTGTAACTGCTACTGCTATTTCATACCAAGGAAAGCTT	791	Qy 479 AGGGAGAAAGCACCCAGGTAAAGTCATGCTGTTTGGAGATGTTAAAAGTGTACACTCA 551 Db 559 GGGGAAAGGACTCAAGGAAAATGGTGTGATCAGGAAATTAATGATCATG 658
Db	839	CTGATGAATTCCTGGCATGCCACCGCCATCTCTCTGGCTGATAGGGAAACTA	898	Qy 432 GAGCTCCGACTAGTTACCTGAACTTCTAGGAAACTTCTGGTAAACTGTTGATGTT 491 Db 539 GAGGCCCTCACGTCACACTCGGGACACCGAGGGCCAGAACGATCAACGATAC 598
Qy	792	CAACATTAGAAATGAGTGTACTCATGACATTAACTTAATTTTAAAGAAAGGAGAT	851	Qy 552 GCTTTCGACTAGTTACCTGAACTTCTAGGAAACTTCTGGTAAACTGTTGAGTT 611 Db 659 GTTTGGCTCTGGTAACTCATCTTAAAGGCAAAATGGAGAACCCCTTGAAGTC 718
Db	899	CAGCAGCTGGAAATGAACTCACCCACGATATCACCAAGTGTGACCATC	958	Qy 612 AAAGATACTGAGAGGAGATTTCTAGTCAAGGCTTAACTATTGCTGATGTT 671 Db 719 ANGGACACCGAGGAAGGGACTTCAGCTCAGGACTGACCTGAGGGTGCCTATG 778
Qy	852	CGTCGTAGCGCTTCCTGCACTCTGCAACCTGCAAACTTAATGTTACCTACGACTTAAA	911	Qy 672 ATGAAAGACTGGTAAACCTGAACTTCTAGGAAATGGTAAAGTGTGAAAGCTT 731 Db 779 ATGAGGGTTAGGCTTAACTCTGAACTTCTAGGAAATGGTAAAGTGTGAAAGCTT 738
Db	959	AGAAAGTCGTGCAAGCTTACCTTACCCAAACTACTGTGCACTATGTCAG	1018	RESULT 14 I04196 Locus 104196 bp DNA linear PAT 02-DEC-1994
Qy	912	TCTGTGTTAGGCCAGTTAGGTTTACCAAACTGTTTCAACGTTGCTTGTAGTGT	971	Qy 104196
Db	1019	AGCGTCCTGGGCTCACTGGCATCTACTGGTCACTTAACTGTTCTCCGG	1078	Qy 1032 GATGAAAGGAGGACTCCAGGCGCCGCGCTATGTTCAATGAGCATT 1091 Db 1139 GACGAGAAAGGACTCAACTCCAGGCTGTTAGGGCCTACCCATGTCATC 1198
Qy	972	GTACTGAGAGGAGGTCATTAATTTGACTAAAGCTGTTCAAAAGCCCTTAACATT	1031	Qy 1092 CCACCCAGAACTTAAATAACCATGTTTCTGATGTCAGGAAACACTAA 1151 Db 1079 GTCACAGGAGGAGGACCCCTGAAGCTCTCAAGGCCGTCATAAGGCTGTCACCATC 1138
Qy	1032	GATGAAAGGAGGACTCCAGGCGCCGCGCTATGTTCAATGAGCATT 1091	Db 1199 CCCCGAGGTCAACTTCACAAACCCCTGTGCTCTTAATGTTGAACAATACCCAG 1258	
Db	1139	GACGAGAAAGGACTCAACTCCAGGCTGTTAGGGCCTACCCATGTCATC 1198	Qy 1152 AGCCCATGTTTGTGTTAGGGTAAAGGTGTCACCCAACTCAGA 1192 Db 1259 TCTCCCTCTCATGGAAAAGTGTGAATCCACCCACCCAAA 1299	
Qy	1152	CCACCCAGAACTTAAATAACCATGTTTCTGATGTCAGGAAACACTAA 1151	Qy 1259 TCTCCCTCTCATGGAAAAGTGTGAATCCACCCACCCAAA 1299	
Db	1259	TCTCCCTCTCATGGAAAAGTGTGAATCCACCCACCCAAA 1299	Qy 132 TTAATGAAGTATTAGGTTAACCGCTACTGTTACATCAGACTGTAAGAGCTT 791	

839	CTGATGAAATACTGGCAATGCCAACGGCATTCTTCTCCGATGAGGGAACTAA	898
792	CRACATTAGAACATGAGTTGACTCATGACATTATTAACATAATTCTTGTAGAACAGGAGAT	851
899	CGACCTCTGGAAAATGAGCTACCCACGATATCATACCAAGTAAATGAGAC	958
852	CGTGTAGCGCTCTCGACCTGCCAAGHTAAGTATCACCGTACTTACGACTTAA	911
959	AGAAAGCTGGCAGCTTACATTTACCCAAACTGTCCATTACTGAACTTATGTCGA	1018
912	TCGTGTTTAGGCGAGTAGGTTTACCAAAAGTTTCTACGGTCCGATTGACTGGT	971
1019	AGCTGTTCTGGCAACTGGGATCACTAAGGCTTCAAGGCTTCAAGCAATGGGCTGACCTCTCGGG	1078
972	GTTACTGAGAACGCTCCATTAAATTSAGTAAGGTTCTACAAAGCGGTAACTATT	1031
1079	GTCACAGGGGGCACCCCTGAAGCTCTCCAGGGCGTGCATAAGGCTGTCGACATC	1138
1032	GATGAAAGGTACCGAGCGCCGGCTAATGTTCTGGAGGTATCCATGACCTAA	1091
1139	GAGGAGAAGGGACTGAAGCTCTGGGCCATGTTAGGGCCATACCCATGCTATC	1198
1092	CCACCAAGGTTAAATTAAACCATTCGTTTCTGTATGAGGAGACACTAA	1151
1199	CCCCCGAGGTCAAGTCACAAACCTTTGTTCTGTCTTATGATGACAAATACCAAG	1258
1152	AGCCCATGTTGTTGTTGAGCTTGTCAACCCAAACTCAGAA	1192
1259	TCTCCGCTCTCATGGAAAATGTTGATTCACCCAA	1299

RESULT 15
104272 LOCUS
DETERMINATION Sequence 1 from Patient
104272
PAT 0139383
DNA linear
PAT 02-DEC-1994

UNCLASSIFIED
REFERENCE 1 (bases 1 to 1434)
AUTHORS Russell, P.R.
TITLE Method for expressing foreign genes in *schizosaccharomyces pombe*
and the use in therapeutic formulations of the products. DNA
constructs and transformant strains of *schizosaccharomyces pombe*
usable in such method and their Preparation
Patent: EP 019383-A1 102 MAY 1985;
JOURNAL

source	1. 1434	/organism="unknown"	293 t
BASE COUNT	369 a	432 C	340 g
ORIGIN			

Query Watch score 429; DB 6; Length 1434;

be Metabolos 7.11: *Conservative* 0: Mismatches 420: Index 0: Gaps

卷之三

卷之三

卷之三

BB 239 CTGGCACACCAACAGCACCAATCTTCTCCCCAGTGAGCATCGCTACAGCC 298

Gu 192 T¹TCGCCATGTTGAGTTAGGTACTAAAGCCGATAACCCGATGACCAAGATTAGAAGCTTA 251

Qy	252	AACTTTAATTGACCGAAATTCAGGCCAAATTACCGAGGGTTTCAGAGTTGTTG	311
Db	359	AAATTCAACCTCAGGAAATTCGGAGCTAGATCCAGAGGGCTTCAGAAACTCCCTC	418
Qy	312	AGAACTTGTAACTGTATTCTCAATTGCAATTAACTGTAACTATAGTTTTTG	371
Db	419	CGTACCTCAACCGCCAGACGCCAGCTTACCCGGZATGGCTGTCCTC	478
Qy	372	TCTGAGGTTTAAATTTGGTGCACAAATTCTAGCAAGAACCTATATAGT	431
Db	479	AGCGGGCCCTGAAGCTGAAGTAACTTTGGAGATGTAAAGTGTACACTCA	538
Qy	432	GAGGCTTTACCGTTAATTGGTGTAACTGTAGGAAGCTAAAGGAAATTAAATGATTAT	491
Db	539	GAAGGCTTCAGTCAACTTCGGAGACCCGAAAGAACAGATAACGGATTAC	598
Qy	492	GTTGAGAAAGGACCCAGGTAAGATGTTGACCTAGTTAAAGATTAGATCGTATACC	551
Db	599	GTCGAGAAGGGTACTCAGGGAAATGTTGTTGAACTTGTGAGCTTCAAGAGACACA	658
Qy	552	GTCCTCGCACTAGTAACTATTTTCAAGGTTAAAGTGGAAACGCTTTCGAGGT	611
Db	659	GTTTGGTCGGTAATTACATCTCTTAAGGCAATGGAGAGACCCCTTGAGTC	718
Qy	612	AAAGATACTGAGAGGAAGATTTCATGTTGATCAAGTTACTACTGTCAAACCTCCATATG	671
Db	719	AACGACACCAGGAAAGGACTTCACGGCACTTGACCGTGGACCGTGAAGCTGGCTATG	778
Qy	672	ATGAAAAGACTGGTAGTTCAATATICAACATTGCAAAATTAACTTAAAGTCCTT	731
Db	779	ATGAAAGCGTTAGGCACTGTTAACATCCAGCACTGTAAAGCTGTCAGCTGGCTGT	838
Qy	732	TTAATGAGTATTAGGTAACTGCTACTGCTATTTTTACAGGAAAGTAACTGTT	791
Db	839	CTGATGAAATACCTGGCAATGCCACGCCATCTCTCCIGCCTGTAGGGAAACTA	898
Qy	792	CAACATTAGAGAAATGAGTGTGACTCATGACATTAACTAAATTAGAGAACGAGGT	851
Db	899	CAGGACCTGGAAATAGACTCAACCCAGCATACATCCAAATTCCTGAAATAGAAC	958
Qy	852	CGTGTAGGGCTTCCTGCACCTGCCAAAGTTAAAGTATCACGGGTACTTACGACTAAAA	911
Db	959	AGAAAGGTCTGGCAGCTTACATTTACCCAAACTGTCCTATTACTGAACTTATGTCAG	101
Qy	912	TCTGTATTAGGGCACTGAGTATTACCCAAAGTTCTAACTGGGCAATTGAGTGT	971
Db	1019	AGCGCTCTGGCTCACTGGCATACTAAGGCTCTCAGCAATGGGTGACCTCTCCGGG	107
Qy	972	GTAACTGAGAAAGCTCCATTAAATTTAGTAAGTAAAGCTGTTCAAAAGCTTAACTATT	103
Db	1079	GTCTCACAGGAGGTCAACAAACCCCTTGTGTCACATGAACTGACAAATACCAAG	113
Qy	1032	GATGAAAAGGGTACCGAGGCCCTATGTTCTGGAAAGCTTCAATGAGCATT	109
Db	1139	GACGAGAAAGGGACTGAAACTGCTGGGGCCATTTTGTAGGSCCATACCCATGTCPT	119
Qy	1092	CCACAGGTAAATTAAATAAACCTCAGTTCTGATGTCAGGAGAACACTAA	115
Db	1199	CCCCCGAGGTCAACAAACCCCTTGTGTCACATGAACTGACAAATACCAAG	125
Qy	1152	AGCCATTGTTTATGGTAAAGGTTGTCACCCAACTCGAA	1192
Db	1259	TCTCCCTCTCATGGAAAAGTGTGAATCCACCCAAA	1299

Search completed: December 7, 2002, 01:14:57
Job time : 2698.5 secs